

```
import numpy

import pandas as pd

# Step 2: Load the dataset
df = pd.read_csv("food_coded.csv")

# Step 3: Data Cleaning

# Fill missing values with mean, median, and mode
for column in df.columns:
    if df[column].dtype in ['int64', 'float64']:
        df[column].fillna(df[column].mean(), inplace=True)
    else:
        df[column].fillna(df[column].mode()[0], inplace=True)

# Remove duplicate rows
df.drop_duplicates(inplace=True)

# Remove duplicate columns
df_transposed = df.T
df_transposed.drop_duplicates(inplace=True)
df = df_transposed.T

# Step 4: Save the cleaned dataset
df.to_csv("cleaned_dataset.csv", index=False)
```

```
# @title  
df.head()
```



	GPA	Gender	breakfast	calories_chicken	calories_day	calories_scone	calories_veg
0	2.4	2	1	430	3.028302	315.0	420.0
1	3.654	1	1	610	3.0	420.0	420.0
2	3.3	1	1	720	4.0	420.0	420.0
3	3.2	1	1	430	3.0	420.0	420.0
4	3.5	1	1	720	2.0	420.0	420.0

5 rows × 61 columns

Start coding or [generate](#) with AI.